Dimension Reduction:

April 13 2021

JL Transform:

- Linear Transform — E.g. Ax (linear transform)
  - Mx. mult
- Preserve pairwise L_2 distances

Set of points $X = \{x_1 \dots x_n\}$ in $\mathbb{R}^D$ ← original dimension

Transform $S : \mathbb{R}^D \longrightarrow \mathbb{R}^K$ ← final dimension (after the reduction)

$$x_i \longrightarrow Sx_i$$

JL Lemma/Thm :

Let $\varepsilon \in (0, 1/2)$, $X = \{x_1 \dots x_n\}$

There exists a map $S : \mathbb{R}^d \to \mathbb{R}^k$

with $\boxed{K = O\left(\frac{\log n}{\varepsilon^2}\right)}$

$\forall\, i,j$ in $\{1 \dots n\}$

$$(1-\varepsilon)\|x_i - x_j\|^2 \leq \|Sx_i - Sx_j\|^2 \leq (1+\varepsilon)\|x_i - x_j\|^2$$

$$P(|x - \hat{x}| \leq \varepsilon \alpha) \geq 1 - \delta$$

$$X_i \longrightarrow Sx_i$$

$$(1 \pm \varepsilon)\|x_i - x_j\|^2$$

- Log factor reduction

10 billion $= 10^{10} \simeq 2^{30}$ $\qquad K = O\left(\frac{30}{\varepsilon^2}\right)$

Construction:

Let $\underset{(K \times D)}{A} = \begin{bmatrix} & & \\ & \circ & \\ & & \end{bmatrix}$

$\sim \mathcal{N}(0,1)$ i.i.d $\rightarrow$ distributed as...

Transformation matrix $S \triangleq \frac{1}{\sqrt{K}} A$

$x \in \mathbb{R}^D$ mapped to Sx

Proof sketch:

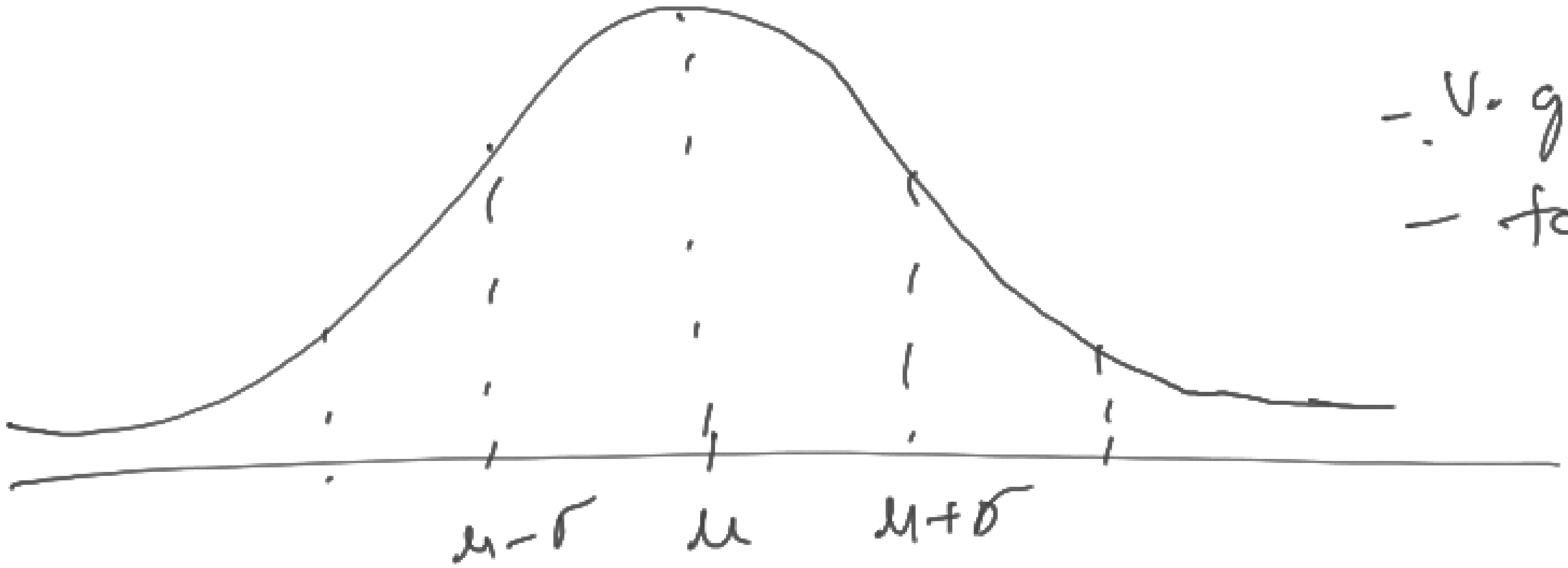
Gaussian R.V.s:

Cont. R.V.

Denoted $\mathcal{N}(\mu, \sigma^2)$

mean variance

PDF $f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}$



- V. good concentration
- tapers off exponentially $e^{-\frac{y^2}{2\sigma^2}}$

Properties :

P1. $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ & independent

$$X_1 + X_2 \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

P2. $X \sim \mathcal{N}(\mu, \sigma^2)$

$$aX \sim \mathcal{N}(a\mu, a^2\sigma^2)$$

P3. $X_i \sim \mathcal{N}(0, 1)$ → "Standard normal"

$a = [a_1 \dots a_n]^T$

$$\sum a_i X_i \sim \mathcal{N}(0, \sum a_i^2)$$

$$\mathcal{N}(0, \|a\|^2)$$

L_2 norm / Euclidean length of "a"

.... to J L Transform.

Fact 1 : Let $X \in \mathbb{R}^D$

$$\underset{(k \times D)}{A} = \begin{bmatrix} — A_1 — \\ — A_2 — \\ \vdots \\ — A_K — \end{bmatrix}$$

$$A \; X = \begin{bmatrix} — A_1 — \\ \vdots \\ — A_K — \end{bmatrix} X = \begin{bmatrix} \langle A_1, X \rangle \\ \langle A_2, X \rangle \\ \vdots \\ \langle A_K, X \rangle \end{bmatrix}$$

$$\langle A_1, X \rangle = \sum_{i=1}^{D} A_{1i} X_i \sim N(0,1)$$

JL Transform : $Sx = \frac{1}{\sqrt{K}} A x = \frac{1}{\sqrt{K}} \begin{bmatrix} \langle A_1, x \rangle \\ \langle A_2, x \rangle \\ \vdots \\ \langle A_K, x \rangle \end{bmatrix}$

$$\sim \begin{bmatrix} N\left(0, \frac{\|x\|^2}{K}\right) \\ N\left(0, \frac{\|x\|^2}{K}\right) \\ \vdots \\ N\left(0, \frac{\|x\|^2}{K}\right) \end{bmatrix} \quad K\text{-length}$$

Fact 2: For any $Y \in \mathbb{R}^D$

$$E\left[\|SY\|^2\right]$$

$$= E\left[\sum_{i=1}^{K} \frac{1}{K} \langle A_i, Y \rangle^2\right]$$

$$= \sum_{i=1}^{K} \frac{1}{K} E\left[\langle A_i, Y \rangle^2\right]$$

$$= E\left[\langle A_1, Y \rangle^2\right]$$

$$= Var(\langle A_1, Y \rangle)$$

$$\boxed{E\left[\|SY\|^2\right] = \|Y\|^2}$$

Sx linear

$$\|Sx_i - Sx_j\|^2$$

$$= \|S\underbrace{(x_i - x_j)}\|^2$$

$$= \|SY\|^2$$

$Y = x_i - x_j$

$$\langle A_1, Y \rangle = \sum_{i=1}^{D} A_{1i} Y_i$$

$$\boxed{Var(z) = E[z^2] - (\overbrace{E[z]}^{0})^2}$$

$$E\left[\langle A_1, Y \rangle\right] = 0$$

Expected value of the pairwise distances after transformation

$$E\left[\|Sx_1 - Sx_2\|^2\right] = E\left[\|S(x_1 - x_2)\|^2\right]$$

$$= E\left[\|SY\|^2\right] \qquad Y = x_1 - x_2 \in \mathbb{R}^D$$

$$= \|Y\|^2 \qquad \text{(by Fact 3)}$$

$$\boxed{E\left[\|Sx_1 - Sx_2\|^2\right] = \|x_1 - x_2\|^2}$$

w.h.p

$1/n^c$

Gives us what we want in expectation.

To do: take to high prob. result

Lemma 1 (Concentration bound)

Let $U_1 \dots U_K$ be i.i.d. $\mathcal{N}(0, \sigma^2)$

Let $Z = \sum_{i=1}^{K} U_i^2$

$E[Z] = K\sigma^2$

Then for $\varepsilon \in (0, 1/2)$, some constant c

$$\boxed{\Pr\left[\,|Z - E[Z]| \geq \varepsilon E[Z]\right] \leq e^{-\frac{K\varepsilon^2}{c}}}$$

χ^2 - (Chi-squared) distribution

$E[Z] = \sum_{i=1}^{K} E([U_i^2])$

$K = O\left(\frac{\log n}{\varepsilon^2}\right)$

Going back to JL Transform..

Think about this... (next lecture)